Data Integration: Semantic Approach

FORCEnet Engineering Conference

November 15, 2005

Dr. David J. Roberts
619.540.1814
roberts@spawar.navy.mil

Outline

► What is the Problem?

► What is Needed?

► Details

► Examples

► Long Range Goals

What is the Problem?

► Floods of "good" data

► Hundreds of "good" sources

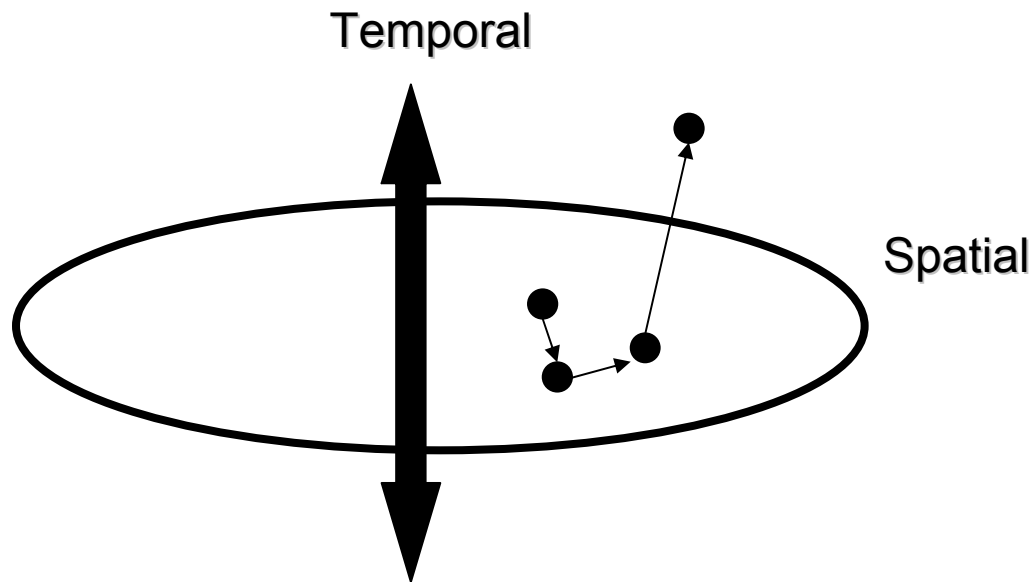► Thousands of "good" people

Result: Many "good" solutions, but…

► Often the pedigree of data is lost

► Contexts rapidly change and solutions are no longer appropriate

► Opportunities are missed in the "noise" of data

Data lies in multiple dimensions of interest (views), in dissimilar temporal and spatial domains, in a dynamic environment.

Temporal

Spatial

All the parts are moving!

What is Needed?

► A common vocabulary for the exchange of data at a semantic level (Machine to Machine (m2m))

► Establish data context using common ontologies

– Domain/Attribute classification, i.e. FORCEnet Consolidated Compliance Checklist (FCCC)

–

– Relationship/Pattern structure, i.e. attached to, welded on, part of

► Agile decision support tools to assist subject matter experts

Result: Agreed upon "best" solutions…

Better ways to manage data… A vendor independent standard to separate style, content, context, and presentation:

► eXtensible Markup Language (XML, "son of SGML, 1981"), 1996, was a good idea… data and commonly defined name spaces

► Resource Description Framework (RDF), 1999, was better… publishes the data, name space, and schema (60K published news feeds) Note: The Virtual Knowledge Repository (VKR) integrates 681 technology news feeds and several databases

► Ontology Web Language (OWL), 2004, DARPA… Supports Description Logic (DL), computational reasoning, and Service Oriented Architecture (OWL-S)

Details

"A common environment for the exchange of data at a semantic level (m2m)"

– Automated data exchange between systems

– Vocabularies that can dynamically adapt to new context and needs

– Underlying data structures are unaffected

– Visualization tools to allow non-semantic experts to see the associations of information as the context changes to understand options and opportunities… adding to mission success
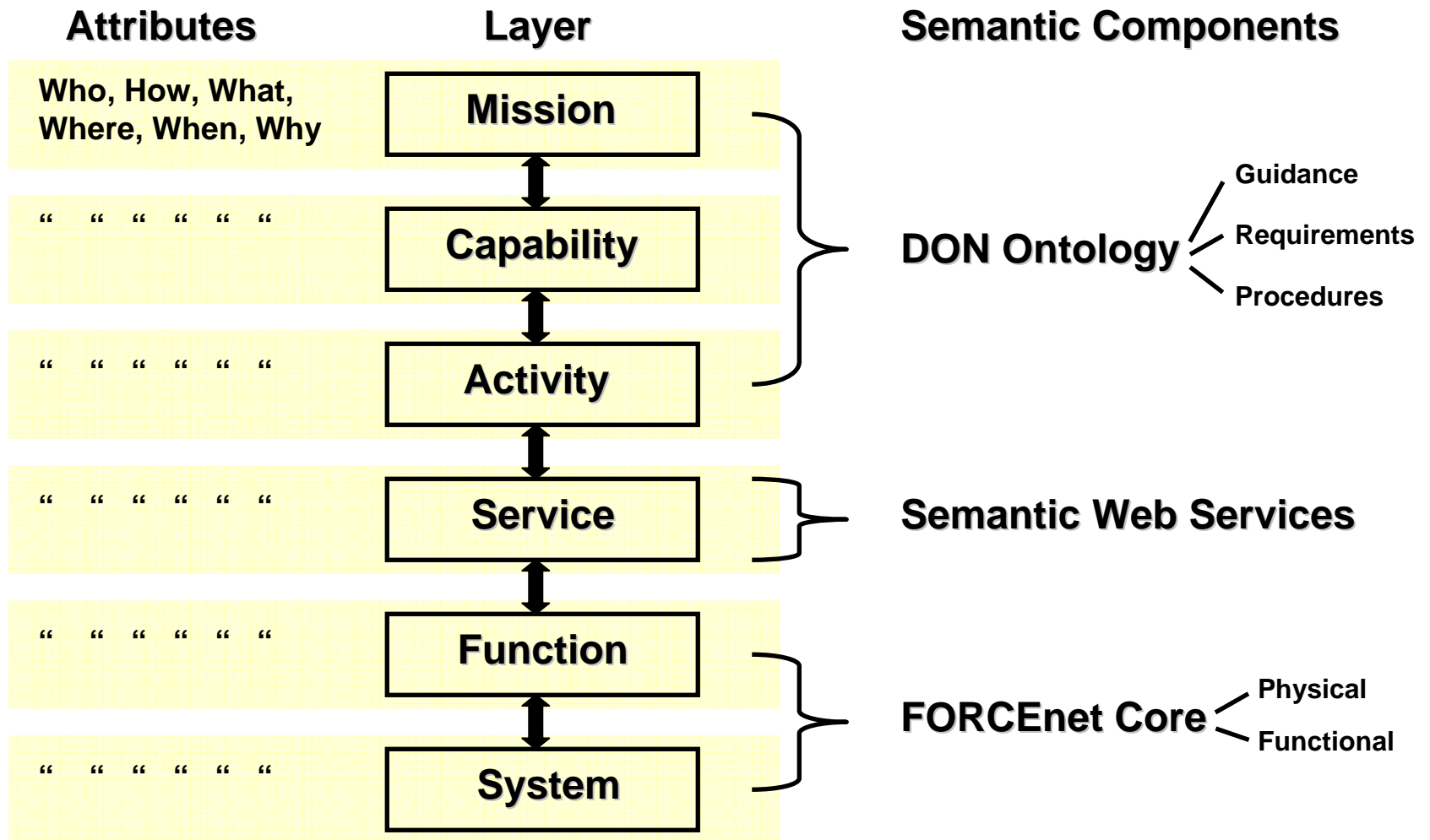
Examples of Semantic Implementations

► FORCEnet Implementation Base Line (FIBL), Spiral 2

- – Semantic "tagging" of requirements
- – Semantic association of data (FCCC taxonomy)
- – Semantic mapping of program profiles

► FORCEnet Implementation Tool Suite (FITS)

- – DON profile classification (ontology)
- – FORCEnet Core (FnC) standardized vocabulary
- – Fn publication mechanism (UDDI based)

► Policy Visualization Tool (demo)

- – Re-uses components of prior Navy efforts
- – Ontology driven document model

The Semantic Layer Cake

Attributes	Layer	Semantic Components
Who, How, What, Where, When, Why	**Mission**	
" " " " " "	**Capability**	**DON Ontology** — Guidance, Requirements, Procedures
" " " " " "	**Activity**	
" " " " " "	**Service**	**Semantic Web Services**
" " " " " "	**Function**	**FORCEnet Core** — Physical, Functional
" " " " " "	**System**	

The Service Layer

Machine to Machine Information Traffic

Resource (description)

FORCEnet Core

Profile (context)

DON Ontology

UDDI Services

WSDL (ebXML)

Other Services

Grounding Ontology

Other Ontologies
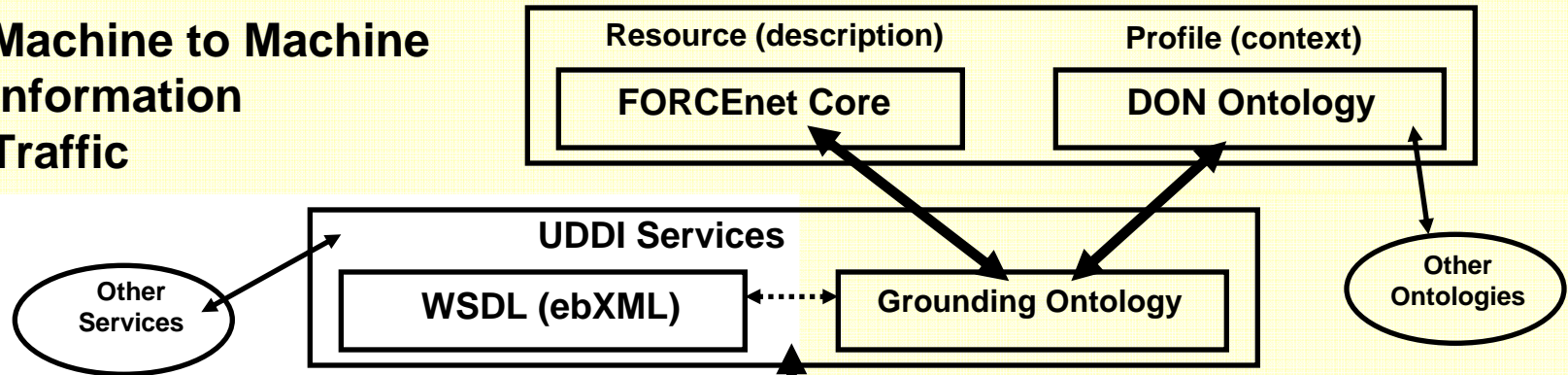
Traditional Information Traffic

Service/Message

Security	**Message**	**Content**	**Transport**

Transaction	**Reliability**	**SOAP**	**HTTP**
Coordination	**Routing**	**XML**	**SMTP**
Context	**Address**	**XSLT**	**Other**
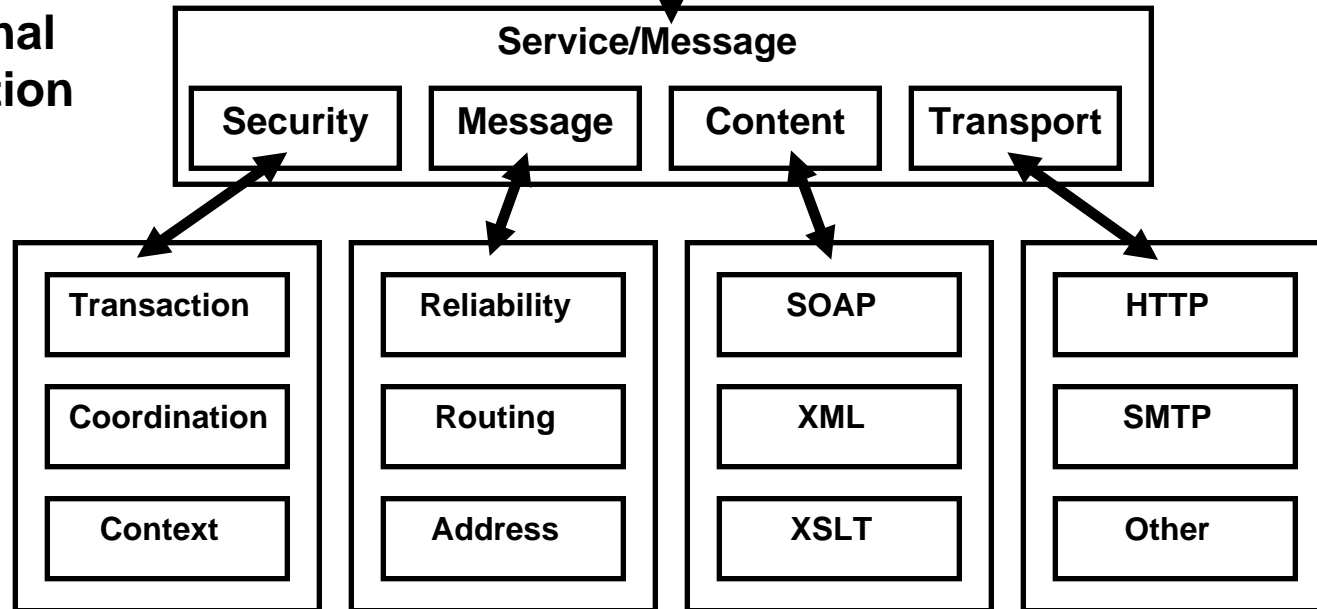
Continuing Goals

► Provide an architectural framework for fielding new data management, correlation and visualization tools…

   Example: The current FORCEnet Core vocabulary

► Integrate emerging semantic technologies into the Navy's net-centric warfare operations construct…

   Example: The current FIRES/Edison/TXC/VKR integration

► Implement semantic-driven knowledge acquisition and discovery capabilities…

   Example: The current FORCEnet Implementation Tool Suite (FITS) semantic environment

Questions ?